

The REPAL 2016 Annual Meeting featured a discussion about the Data Access and Transparency Initiative (DA-RT) initiative. The discussion was led by Evan Lieberman of MIT and attended in plenary session by all participants of the conference. Below is a summary of Evan Lieberman's remarks and the following discussion.

INTRODUCTORY REMARKS BY EVAN LIEBERMAN

Evan Lieberman explained that journal editors are taking specific steps to formulate standards for sharing data and methods for analyzing that data in addition to what is appearing in the articles and appendices and those standards will have implications for scholars about how we do research and value of different types of work, especially qualitative work as well as the way in which we carry out fieldwork.

Concerns that have been raised focus on the trade-off between transparency and privacy and security of subjects. In addition, there has been a polarization that resembles the Perestroika one and that suggests that there are scholars who perceive this process as disrespectful of diverse research traditions.

What are the key issues in the debate?

- How much original data should researchers be required to share with the public and on what terms? Should the norms of DA-RT apply equality to quantitative and qualitative?
- The current norms establish clear requirements for quantitative data, which involve making dataset public as well as the codes that generate the results for others to re-run the analysis. These norms serve to guarantee the reproducibility of results, robustness of results to model specification, data as a public good that should be shared and therefore increases care taken of data analysis for publication.
- Should the qualitative data be shared in the same way? This could involve sharing transcripts from interview, focus group discussions, and archival material. It would provide a public good and would avoid cherry picking of confirming evidence.

Increasing transparency with the data would increase the replicability and reliability of what is published and will garner more respect to qualitative research including the possibility of replication.

What are the objections raised so far? Difficulties for protecting subjects' anonymity (security, reputation).

Possible solution: If qualitative scholars have these concerns they should not share these particular transcripts.

Also, it is important to include the context to the transcripts. Even if transcripts are released, replicability is not possible without context, which shapes the interpretation of transcripts. One possible solution is to annotate transcripts with cultural context and appropriate interpretation.

As suggested by Kathleen Thelen of MIT, it should be an “opt in” rather than “opt out” strategy. That is, at least initially, scholars should be encouraged to share data, but the decision to not share should not be subject to review.

There is a [qualitative transparency deliberation blog](#) (put together by Tim Buthe of Duke University and Alan Jacobs of the University of British Columbia). Anybody interested should go to it to follow the debate.

A lot of qualitative research is based on historical observation and interviews that can be shared, and two options for greater transparency would be to the active citation strategy proposed by Andrew Moravcsik of Princeton University and Evan Lieberman’s own piece on replication databases in the Annual Review of Political Science.

An important criticism of the DART-related proposals is that it will generate disadvantages for younger scholars. But this can be solved by setting good frameworks for data collection and good training for grad students and by rewarding those who do this rather than punishing those who do not. Especially, because sharing data increases citations and use by others of data that otherwise goes unused.

DISCUSSION

MATTHEW AMENGUAL:

How you reward without punishing the other paper that is equal but cannot publicize the transcripts in a journal decision?

STEPHEN KAPLAN:

Will there be intermediate situations when only part of the transcript can be released?

EVAN LIEBERMAN:

Reward without punishing is possible, it takes a lot of effort to prepare data to be shared with the world and that effort should be rewarded. Dialogue with anthropology?

YURI KASAHARA:

I am based in Norway and the government of Norway forces scholars to erase transcript after one year by national legislation. How are scholars in those countries going to operate with new DART rules to publish in US journals?

EVAN LIEBERMAN:

Openness in government should bring norms of openness to expand and may change rules of privacy.

Political Science is in the infancy of qualitative work. Many times, the work does not have explicit LINKS between data and inference. We need more data on how our observations work for us.

CARLOS RUFIN:

I am at a Business school and business journals publish qualitative pieces with rigorous methods for analyzing qualitative data (use of software such as INVIVO). Will this happen in Political Science?

EVAN LIEBERMAN:

KKV applies quantitative inference methods to qualitative research and it was widely criticized but very successful because it was clear and transparent. Jim Scott changed the field with his qualitative work that may be hard to reproduce, but for those who don't have extraordinary skill such as Jim Scott and those who changed the field without a methodological reproducibility, it may be good to have norms and methods to guide us.

LUIS SCHIUMERINI:

What do you think of pre-registration requirements?

EVAN LIEBERMAN:

It makes sense for experiments but pre-registration for qualitative methodology is more difficult and does not allow inductive theory building.

The publication only of positive outcomes and hiding of null results is what brought the requirement of pre-registration of your hypothesis to avoid your hypothesis being a post-hoc finding explanations. This is right for experiments, but maybe too much time in Political Science is spent on causal questions while ignoring important descriptive questions. In biomedicine, there are many descriptive articles or articles patterns of disease with correlates, but not causal inference. World is changing and we should be observing and describing it and measuring in ways that are illuminating, but there are not enough incentives for people to do it.

There is now a growing movement for pre--registering EVERYTHING including observational studies, but the crucial practice should be to be honest about the ordering of theoretical building.

BRIAN PALMER-RUBIN:

The context involves the choice of interviews and context of how the interview happened, which often tells me more about evidence than the transcript of the interview itself and it is thus more important than the transcript. That is, who wasn't interview is more important than content.

EVAN LIEBERMAN:

This stuff is so important that it should be in the body of the paper. The choice of interviewees criteria should and rejection rate should be in the article.

KATHY HOCHSTETLER:

Quantitative dataset include data that you coded and nobody asks for the primary sources that you have coded to check => how far one has to go and why are demands stronger for qualitative data? What is the extent of the documents and interviews that need to be provided, are thousands of read pages have to be submitted? It seems that there are different standards than for quantitative data that only requires excel dataset.

ISABELLA ALCANIZ:

The public good argument is better than control & accountability. If it is public good, it can be voluntary and bottom up but not for control & accountability, when it should be top down and should be homogenous.

PABLO ANDRADE:

The US standards are imposed in Latam in a twisted form. These standards will come as instruments of repression to researchers who are trying to get into details of political decisions.

VICKY MURILLO:

Who establishes the rules? Should the editor of each journal have the authority? Would there be possibility of appeals?

CANDELARIA GARAY:

Triangulation is crucial for qualitative data and seems to be out of the discussion.

EVAN LIEBERMAN:

TRIANGULATION should be a standard for all research. Who should decide? Standards of making it OPT IN rather than OPT OUT should be explicit. Different journals should have different standards by journal should allow researchers to shop around them and select where to publish and what to read.

There is a repository at Syracuse created by Colin Elman of Syracuse University and Diana Kapiszewski of Georgetown University

Let's try to experiment and allow young scholars to be innovators and create new standards.

Regarding different standards for quantitative and qualitative research. In the former, if information gets lost in the coding, there are others who ask for the coding sheets and revise the dataset. Scholars need to provide rules for quantitative data coding, there needs to be inter-coder reliability and new versions of the dataset emerged out of this process of checking. Hyperlinks should also be used for quantitative data